Microsoft Exchange Server

Chapter 8
Installing Front End/Back End OWA Servers

Deployment Guide

Abstract

Exchange 2000 servers can be deployed in a two-tier, front-end/back-end server topology that enables highly scalable messaging architectures suitable for Internet or Intranet solutions. This chapter introduces the benefits of the front-end/back-end server topology for the implementation of Outlook Web Access (OWA) and discusses related deployment options such as Network Load Balancing and SSL-secured messaging. This chapter also provides front-end/Back End server installation steps and scenario-specific deployment guidelines.

This is a preview edition of the Microsoft Exchange 2000 Server Upgrade Series. The final documentation will be available in November 2000.

© 2000 Microsoft Corporation. All rights reserved.

THIS IS PRELIMINARY DOCUMENTATION. The information contained in this document represents the current view of Microsoft Corporation on the issues discussed as of the date of publication. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information presented after the date of publication.

This document is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS DOCUMENT.

Microsoft, Active Directory, Windows, Windows NT and the Windows logo are registered trademarks of Microsoft Corporation.

Other product or company names mentioned herein may be the trademarks of their respective owners.

Microsoft Corporation • One Microsoft Way • Redmond, WA 98052-6399 • USA

1000

CONTENTS

INTRODUCTION....1
FE/BE Review – Design and Benefits 2
Exchange 2000 as a Multi-tiered Application 3
Horizontal Scalability 3
Namespace Management 4
Who Should Use FE/BE? 5
Trade-Offs 6
Resource Requirements 6
  Solution Sizing Plan 6
  Other Resources 6

SELECTING APPROPRIATE FRONT-END SERVER HARDWARE....7

PREPARING THE ENVIRONMENT....8
Intranet Deployments 8
Internet-Connected Deployments 8
  Front-End Behind Firewall 8
  Front-End in DMZ 9

INSTALLING FRONT-END/BACK-END SERVERS....11
Before Proceeding 11
Promoting a Front-End Server 11
Configuring a Front-End 13
  HTTP-specific Configuration 14
Configuring a Back-end 14
  HTTP-specific Configuration 14

PLANNING AND CONFIGURING NETWORK LOAD BALANCING....17
Namespace Planning 17
  URL for OWA 17
NLB Configuration 17
  Single Network Adapter Configuration 18
  NLB Configuration Parameters 19
Testing NLB 19
More Information 19

CONFIGURING SSL....20
  Install server certificate on every front-end server 20
Steps: 20

OPERATIONAL RECOMMENDATIONS....21
Disabling Front-End Services 21
Setting up a Front-end/Back-end Topology: Checklist 21
Backup 21
  FE Backup Requirements 22

Modifying the Default Exchange Virtual Directories 22

SUMMARY AND NEXT STEPS.....24
Hardware load balancing 24
SSL acceleration 24
More Information 24

INTRODUCTION

This chapter introduces Exchange 2000 Server's added support for front-end/back-end (FE/BE) server topologies. Topics related to FE/BE design, deployment and administration are discussed and should be read in their entirety prior to deploying front-end servers. But before we look at this technology in detail, we need to discuss briefly the evolution of Outlook Web Access.

Outlook Web Access in Exchange 2000 is substantially different from the one introduced in Exchange 5.0. Exchange 5.x Outlook Web Access used ASP pages to communicate with the Exchange server that used CDO 1.2 and MAPI. In this context, Outlook Web Access for Exchange was actually a part of Internet Information Services (IIS). The reliance on MAPI-based access to the Exchange server for accessing the mailbox store limited the effective number of people who could simultaneously use the service.

Unlike previous versions of Outlook Web Access, Outlook Web Access in Exchange 2000 does not use MAPI to communicate with the mailbox store, and ASP pages are no longer used for client access. Clients continue to access the Web Access Component through HTTP; however, the IIS server hosting the component uses the new Web Store to provide access to the user's messaging functions.

Internet Information Services receives Outlook Web Access client requests as a proxy for message traffic between a Web client and an Exchange server. Internet Information Services receives the client requests, looks at the namespace, and passes the client to the appropriate Outlook Web Access scripts registered for that URL. If the server contains the Exchange 2000 database, Outlook Web Access uses a high-speed channel to access the mailbox store. If the server is a front-end server, Outlook Web Access proxies the request to a back-end server by using HTTP.

By the end of this chapter, you will be able to evaluate the appropriateness of a FE/BE topology for your organization, design a scalable Exchange server topology, and integrate the server topology into an Internet or Intranet environment. The end result will be a flexible application architecture that can be scaled rapidly and optimized for your organization's network integration, client access, storage and security requirements.

The following tasks will be covered in this document:

- Preparing the environment (internal or DMZ) for front-end installation
- Installing front-end server(s)
- Promoting servers to front-end
- Configuring NLB

Preview Documentation

- Configuring SSL for OWA

Preview Documentation

The following areas are not covered.

- Front-end servers and clustering
- Creating virtual servers and non-default virtual directories
- Enhanced OWA and firewall security, including implementing IPSec.

An installation job aid is provided to ensure that all information relevant to the front-end server installation, including optional components such as Network Load Balancing, is available and complete.

FE/BE Review – Design and Benefits

The primary advantage of front-end and back-end server architecture is that you can define a single namespace for users to access their mailboxes. The alternative to a single namespace is to tell users the name of the server storing their mailbox, as you do in a single-server scenario. This complicates administration and compromises flexibility because each time your organization changes and you move mailboxes to another server, you must inform the users. With a single namespace, you can add and remove servers and move mailboxes from server to server, and users can still use the same URL. Creating a single namespace also ensures that Outlook Web Access remains scalable as your organization grows.

Front-end and back-end server architecture provides an additional advantage for organizations that use Secure Sockets Layer (SSL) encryption. All encryption and decryption processing can be handled by the front-end servers, which improves network performance by removing processing tasks from the back-end servers.

Configuring front-end and back-end architecture for Outlook Web Access requires an HTTP virtual server located under the HTTP protocol in System Manager. The default virtual mail directory is *exchange*. The *hostname* you provide for the virtual server is important because it determines the URL that users type to access their mailboxes. The hostname, TCP port, and IP address uniquely identify a virtual server.

The FE/BE server topology is inherited from large-scale Internet messaging environments such as Microsoft Commercial Internet System (MCIS) and Microsoft Hot Mail. These systems achieve high scalability by separating the client communication servers from the messaging database servers. They are examples of multi-tiered architectures, meaning that the core application services are separated and dedicated to function-specific servers within the application architecture. Multi-tiered architectures are the basis for many highly scalable Internet services, including transactional e-commerce applications and database-

driven Web sites.

Partitioning services within a server topology improves the scalability of applications in several ways. Services do not have to compete for server resources when run on dedicated hardware. Servers can be optimized for their function within the application architecture. Should a specific service's performance degrade, it is possible to supplement that service by adding dedicated servers or resources. As a result, application performance can be maintained or even enhanced as application loads increase by scaling the supporting server topology accordingly. Additionally, multi-tiered applications can typically be scaled rapidly since it is easier to deploy a single service than replace an application's entire architecture.

A multi-tiered architecture will scale vertically if additional services or tiers can be added to extend the application's capabilities. Applications can also scale horizontally if additional servers can be readily added to supplement an existing core service tier. In either case, the applications which support vertical and horizontal scalability are more flexible and efficient than flat or single server application models.

A multi-tiered architecture may also improve an application's availability. With redundant servers performing identical application tasks, an application only needs to balance or distribute loads across these servers to increase its fault tolerance and, thus, its availability. Network Load Balancing, a feature included in Windows 2000 Advanced Server, makes this possible by distributing client loads across multiple servers based upon availability.

Exchange 2000 as a Multi-tiered Application

Exchange 2000 can be deployed in a multi-tiered server architecture, and its scalability and availability can be significantly enhanced as a result. Exchange 2000 Server's core services of protocol, storage and directory are partitioned and mapped clearly to the server roles of front-end, back-end and domain controller, respectively.

By default, an Exchange 2000 server hosts public and private stores and, in this role, fulfills the duties of a back-end or storage server. An Exchange 2000 front-end is only responsible for handling client communication requests. Freed from the responsibility of managing a messaging database, a front-end server performs protocol handling and proxy functions at high speed. Both front-end and back-end servers query domain controllers (particularly the Global Catalog servers, or GCs) for directory information that will authenticate users and locate their home servers.

Using this multi-tiered architecture, additional Exchange 2000 front-end

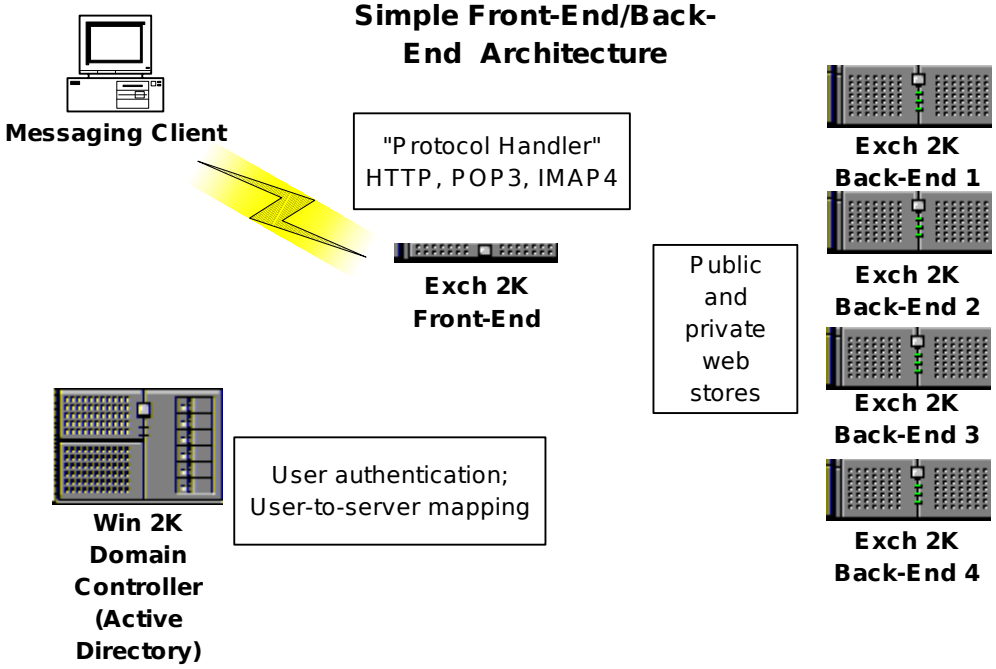
servers may be added as necessary to support increasing client loads. Application availability can also be increased by balancing client requests across multiple front-end servers with NLB. Network load balancing makes sure that user requests are directed to servers that are available, and will help maintain that user's session on a particular server until they disconnect.

Horizontal Scalability

The Exchange 2000 FE/BE topology can be scaled horizontally by adding front-end protocol servers as needed to increase client capacity. Front-end servers are dedicated to communicating with Internet standards-based messaging clients using the HTTP, IMAP4 and POP3 protocols.

As an initial design guideline, deploy a single front-end server for every four back-end servers, as shown in the following graphic.

Diagram 1
Simple Front-End/Back-End Architecture



Additional front-end servers should be considered when client loads cannot be supported by a single front-end server. The Exchange 2000 Resource Kit provides performance monitoring guidelines that will help you determine when a front-end server is reaching its load capacity.

A second front-end server is always recommended for front-end redundancy. When installed in a Network Load Balancing cluster, additional front-end Exchange servers can ensure application availability in the event of front-end server failures.

Namespace Management

A chief benefit of a FE/BE server topology is its support of a unified namespace. A single front-end server name can be used by HTTP, IMAP4 or POP3 clients to access any mailbox or public folder store in the Exchange organization, providing that proper permissions are granted. The messaging client does not need to be configured to connect to a specific back-end server because the front-end queries the Global Catalog server to resolve the user's home server. By proxying and responding to client requests, a front-end server can appear to be the only Exchange server in an organization. The benefit to end users is simplicity. Only one server name is needed for messaging client configuration. In the case of Outlook Web Access, a common URL such as http://email.company.com/exchange can provide access to all mailboxes and public folders, regardless of the actual Exchange server's name.

A single namespace is also possible when more than one front-end server is used. This is achieved by configuring multiple front-end servers in a Network Load Balancing (NLB) cluster.

Who Should Use FE/BE?

If you are using multiple servers in your organization, your best option is using front-end and back-end server architecture to deploy Outlook Web Access. With this type of topology, the front-end server sends HTTP requests to a back-end server running Outlook Web Access. First, the front-end server checks Active Directory to determine which back-end server should receive the request.

The FEBE topology is suited for organizations of all sizes that wish to realize the following scalability, availability and operational improvements:

- Ability to scale Exchange 2000 Server's HTTP (OWA), IMAP4 and POP3 services based upon messaging client requirements.
- Ability to access any mailbox in an Exchange organization with a single server name or with a single URL (in the case of Outlook Web Access).
- Ability to off-load encryption/decryption overhead associated with secure SSL sessions.
- Increased availability of an OWA application when combined with Network Load Balancing
- Enhanced security configurations for a firewall or DMZ environment.

Organizations that use MAPI clients exclusively should not consider a FE/BE server topology since the front-ends only proxy HTTP, IMAP4 and POP3 protocols. MAPI clients connect directly to back-end servers hosting

private and public Web Stores. In cases where MAPI clients request information that cannot be provided by the Exchange server, a re-direct is performed to the appropriate server elsewhere in the organization. This process differs from front-end functionality in that the original client request is referred—and not proxied—by the server. MAPI referral is supported by all Exchange servers, but this feature has no scalability design advantage. Its benefit is that you can move users between servers and their MAPI clients will re-configure themselves automatically.

It should also be noted that configuring an Exchange server to host the front-end processing does not prevent that server from hosting public or private stores. Stores may be accessed on front-end servers just as on back-end servers. As a practice, this is not recommended as it could compromise user functionality and security, depending on the deployment scenario.

Trade-Offs

While many organizations can benefit from a FE/BE topology, the following technical and operational realities need to be evaluated prior to deployment.

A FE/BE architecture will introduce:

- Additional servers
- Network design, integration complexity
- Firewall/security requirements

Network Load Balancing (NLB) may also present integration challenges. NLB is often used in conjunction with FE/BE topologies to distribute client traffic and enhance application availability. This option is easily installed and configured on front-end servers, but administrators will need to understand its functions well in order to integrate it properly in their environment. Additionally, NLB may require reconfiguration of supporting internetworking equipment, such as routers and switches.

Resource Requirements

This chapter is intended for Exchange Administrators with messaging architecture responsibilities. Completion of the installation instructions and job aids provided may also require the assistance of the following individuals:

- **Network Engineer** – to perform non-Active Directory DNS administration, for example, provisioning host names and IP addresses for hosts and NLB clusters.
- **Firewall/Security Engineer** – to perform firewall configuration and communicate security policies regarding server installations in a

perimeter, or DMZ, network.

Solution Sizing Plan

A FE/BE server architecture should be designed in accordance with a server sizing and scaling plan. Developing such a plan is beyond the scope of this chapter, but the plan should specify at a minimum the load capacity estimates, user profiles, service level goals and the appropriate server and storage hardware to meet these requirements.

Other Resources

The following resources may be helpful to you:

- The MS White Paper: “Exchange 2000 Front-End and Back-End Topology,” particularly the section on evaluating the placement of front-end servers and the section on advanced firewall security procedures.
- The Microsoft White Paper, “ Network Load Balancing Technical Overview”
- The Microsoft Exchange 2000 Resource Kit

As a protocol handler, the front-end server process only proxies client requests and does not host any stores. Hence, a single disk or RAID volume for the OS and Exchange 2000 is sufficient for most scenarios. If additional disk resources are available, pagefile performance can be improved by either providing a dedicated drive for its use or by distributing it across multiple drives.

Choose the desired hardware availability features for a front-end server (such as redundant power supplies) by examining the number of front-end servers in your design and your service level requirements. A higher number of front-end servers reduces the need for server availability features, while a design based upon one or two front-end servers should require redundant hardware components. Multiple NICs are recommended when front-ends require Network Load Balancing or adapter fault tolerance.

If multiple front-end servers are planned, consider the space and cost advantages offered by thin, rack-optimized servers, which have less availability features than their full-sized counterparts. This way servers can easily be replaced if the server fails, and you can maintain service levels in a consistent manner.

If significant growth is anticipated, consider developing a scaling plan that answers the following questions:

PREPARING THE ENVIRONMENT

1. How many non-MAPI clients should the solution support?
2. Of those, what is the expected peak user concurrency rate (%)?
3. How many concurrent sessions will each of my front-end servers support based upon observed or anticipated user activity?
4. How many users require concurrent access to a single server name or URL? If this answer is higher than that of the previous question, then determine the number of front-end servers required to support the namespace.

Lastly, consult your server manufacturer for Exchange 2000 server sizing tools or references. Establishing a supportable baseline with a hardware vendor allows you to meet your internal service level agreements better.

This section differentiates and describes front-end server deployments in both Intranet and Internet-connected environments.

Intranet Deployments

An Intranet deployment is one that does not expose Exchange servers to the Internet or an external network for the purpose of providing internal messaging client access. Intranet deployments require no special network integration procedures, but the following server placement guidelines should be observed.

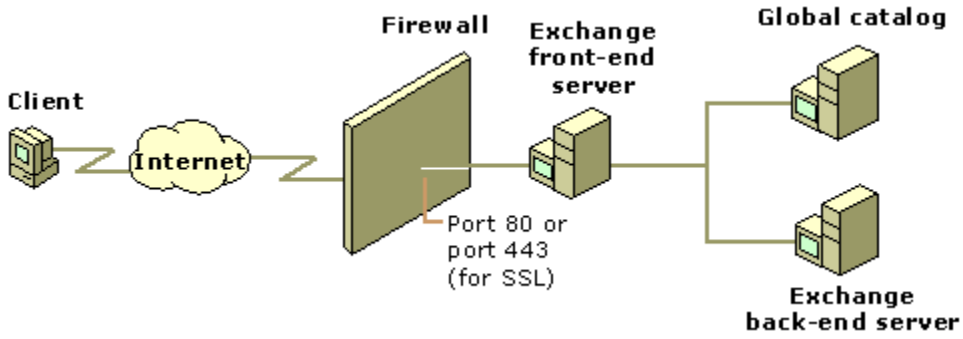
- Do not install front-end servers on domain controllers.
- Place front-end servers on the same segment and site as a Global Catalog server. Front-end servers query GCs to authenticate clients and resolve their requests to back-end servers.
- When possible, place front-end servers on the same segment as back-end servers.
- If you are using Network Load Balancing, all front-end servers participating in the same NLB cluster must reside in the same subnet.

Internet-Connected Deployments

The FE/BE topology provides enhanced security and design flexibility when integrated into an Internet-connected environment. Front-end servers are suitable for deployment in a Demilitarized Zone (DMZ), or perimeter network, since they do not host Web stores or Active Directory databases. This diminishes their value as targets for hackers. Front-ends can also be configured to extend messaging services over a minimum set of software ports, making them ideal for placement behind a firewall or within a DMZ.

Since front-end server placement is flexible, there are two front-end deployment scenarios that should be considered when extending Exchange services to Internet-based clients—front-end behind the firewall and front-end within the DMZ.

Front-End Behind Firewall



If Outlook Web Access clients access the Exchange server over an Internet connection, we recommend protecting the Exchange server with a firewall. The firewall determines whether it is safe to let a file pass through to the network.

Placing a firewall between the client and the front-end server allows OWA clients to communicate with the server using the HTTP protocol and Secure Sockets Layer (SSL) (optional encryption) protocols. These protocols require TCP ports 80 and 443 to be open between the client and the front-end server. The above illustration shows this firewall option.

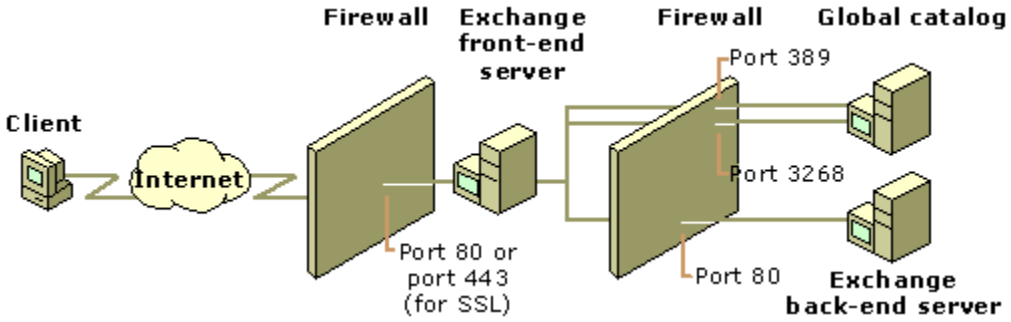
This model may be used by organizations without a DMZ, or whose security policies permit external client access to internal servers.

Enabling this model requires firewall configuration as seen in Table 1.

Table 1. Firewall Configuration: Front-End Behind Firewall

Source	Destination	Service	Protocol and Port
Internet / External Network	Internal / Private Network	HTTP	TCP 80
		HTTPS	TCP 443
		IMAP4	TCP 143
		IMAP4TLS	TCP 993
		POP3	TCP 110
		POP3TLS	TCP 995

Front-End in DMZ



Using two firewalls: one on either side of the front-end server (DMZ). This type of deployment is suited for large organizations. Using this type of firewall configuration provides maximum security. TCP ports 80 and 443 must be open on the firewall between the clients and the front-end server. The TCP ports for network basic input/output system (NetBIOS), remote procedure call (RPC), and Lightweight Directory Access Protocol (LDAP) also must be open, so that the front-end server can send authentication requests to the back-end server and servers hosting Active Directory. The above illustration shows this firewall option with LDAP ports 3268 and 389 open for direct communication between the front-end server and a global catalog server.

Preparing the environment will require external and internal firewall configuration as shown in Table 2. If your organization's security policy does not permit one or more of these services through the internal firewall, it is possible to statically configure services such as DNS and RPC to reduce the number of open ports. For more information on enhanced security procedures, refer to the Microsoft white paper, "The Exchange 2000 Front-End / Back-End Server Topology" (see Other Resources).

Table 2. Front-End Server(s) in DMZ

Source	Destination	Service	Protocol and Port
Internet / External	DMZ Network	HTTP	TCP 80
		HTTPS	TCP 443
		IMAP4	TCP 143
		IMAP4TLS	TCP 993
		POP3	TCP 110
		POP3TLS	TCP 995
		DNS	TCP, UDP 53
		HTTP	TCP 80
		RPC EP Mapper	TCP 135

DMZ Network	Internal / Private Network		
INSTALLING FRONT-END/BACK-END SERVERS		KERBEROS	TCP UDP 88
		LDAP	TCP 389
		NETLOGON	TCP 445
		DSAccess (GC)	TCP 3268
		TCP High Ports	TCP 1024+

Additional ports may be required during front-end server installation. If security policies allow, open all ports between front-end servers and the private network during installation to permit GC communication. Once installation is complete, restrict access as per the table above.

Now lets look at the procedures for installing a front-end/back-end configuration.

Before Proceeding

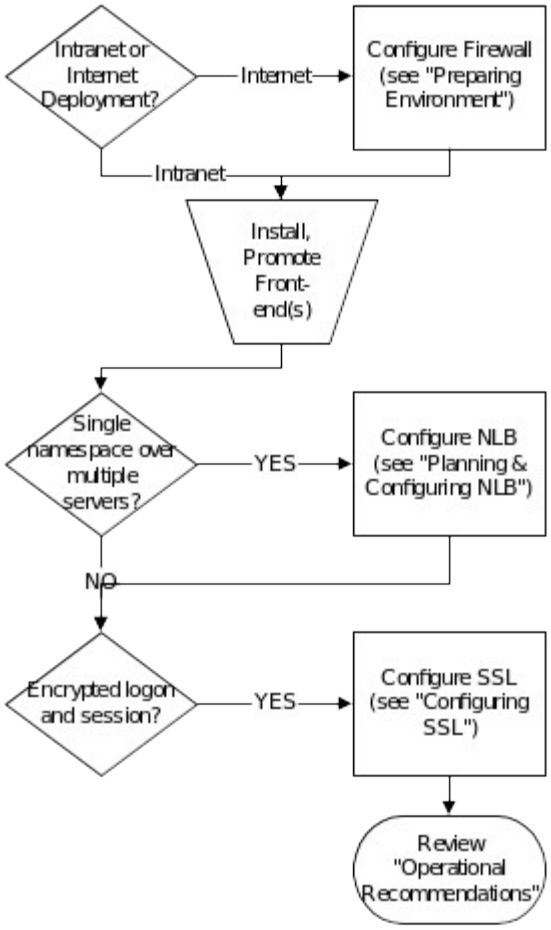
As a best practice, no public or private stores should be hosted on a front-end server. While this is an option that may have appeal to organizations with an interest in preserving server hardware, it is not recommended. This document recommends procedures that puts existing stores at risk and disables access to those stores.

Promoting a Front-End Server

1. After installing the front-end server in an existing Exchange organization, open the Exchange System Manager (ESM), right-click on the front-end server name and select **Properties**.
2. Select **This is a Front End Server**.
3. Click **OK**.
4. Reboot server.

The following flowchart presents the order of installation events for a front-end/ back-end solution.

IMPORTANT: It is assumed that Exchange 2000 servers already exist in the Exchange Organization receiving the front-end server(s).



Configuring a Front-End

A front-end server must be part of the same Exchange organization as the back-end servers. In fact, it is an ordinary Exchange 2000 server until the change. A front-end server should not host any users or public folders.

To set up a front-end server, install a new Exchange 2000 server into the organization.

Then, from the server property page, select the checkbox labeled **This is a front-end server**, and restart the server, or stop and restart the HTTP, POP3, and IMAP4 services.

Before restarting the services, dismount and delete the private and public stores on the front-end server. Failing to do so may make the public store inaccessible from via the front-end server

Not all Exchange Services are required on a front-end, depending on the protocols being exposed. The following list shows the Exchange services required for each protocol; all other Exchange Services can (and should be) be stopped and disabled:

- **HTTP**: No Exchange-specific services required. However, the Windows HTTP service (w3svc) must be running.
- **POP3**: Exchange POP3 (POP3Svc), Exchange Information Store Service (MSExchangeIS), Exchange System Attendant (MSExchangeSA)
- **IMAP4**: Exchange IMAP4 (IMAP4Svc), Exchange Information Store Service (MSExchangeIS), Exchange System Attendant (MSExchangeSA)

To stop and disable services, use the Services Computer Management MMC Snap-in.

By default, IMAP4Svc and POP3Svc are marked has having a dependency on MSExchangeIS. In practice, however, this is strictly not needed when running as a front-end service. Having the store process (without any mailbox or public stores mounted) run on the front-end should not pose any security risk, nor should it affect performance materially. However, it is possible to manually remove this dependency by removing MSExchangeIS from the following registry key:

```
HKEY_LOCAL_MACHINE\
  System\
    CurrentControlSet\
      Services\
       IMAP4SVC | POP3SVC\
         DependOnService
```

Warning: Changes to the Registry can affect the stability of your system. Always back up your computer before changing Registry settings. Microsoft does not support changes to the Registry and all such changes are made at your own risk.

If you are running POP3 and IMAP4 in a configuration where you do not have the RPC ports (135, and 1024+) opened between the front-end server and the back-end servers, you must remove these dependencies and not run the MSExchangeIS service on the front-ends. Refer to the Configuring Firewall section for additional firewall information.

HTTP-specific Configuration
The front-end server's virtual directories and HTTP virtual servers must exactly match those of the back-end. In a default setting, no additional configuration need be done on the front-end—the "exchange" and "public" virtual directories are already matching.

Configuring a Back-end

Back-end servers must be specially configured to support the front-end server. First, they must be servers in the same domain as the front-end. Users may be on any of the servers in the domain. Public folders can be replicated in any way that is desired. Back-end servers can be accessed directly if required, with no effect on the behavior of the front-end/back-end set up.

HTTP-specific Configuration
Extra virtual directories or virtual servers not available on the front-end can be created on the back-end servers to support specific applications.

The main configuration that must be done on the back-end is to create, using the Exchange System Manager, a virtual server that maps to each front-end namespace. In the typical environment, it is expected that all the front-ends would be mapped to a particular name, so as to unify the namespace. The new virtual server on the back-end will be mapped to the single name that represents all the front-ends. This virtual server will respond to the requests from the front-end, and, as such, must be configured to match the configuration on the front-end, including the virtual directories (in particular, "exchange" and "public" in the default configuration).

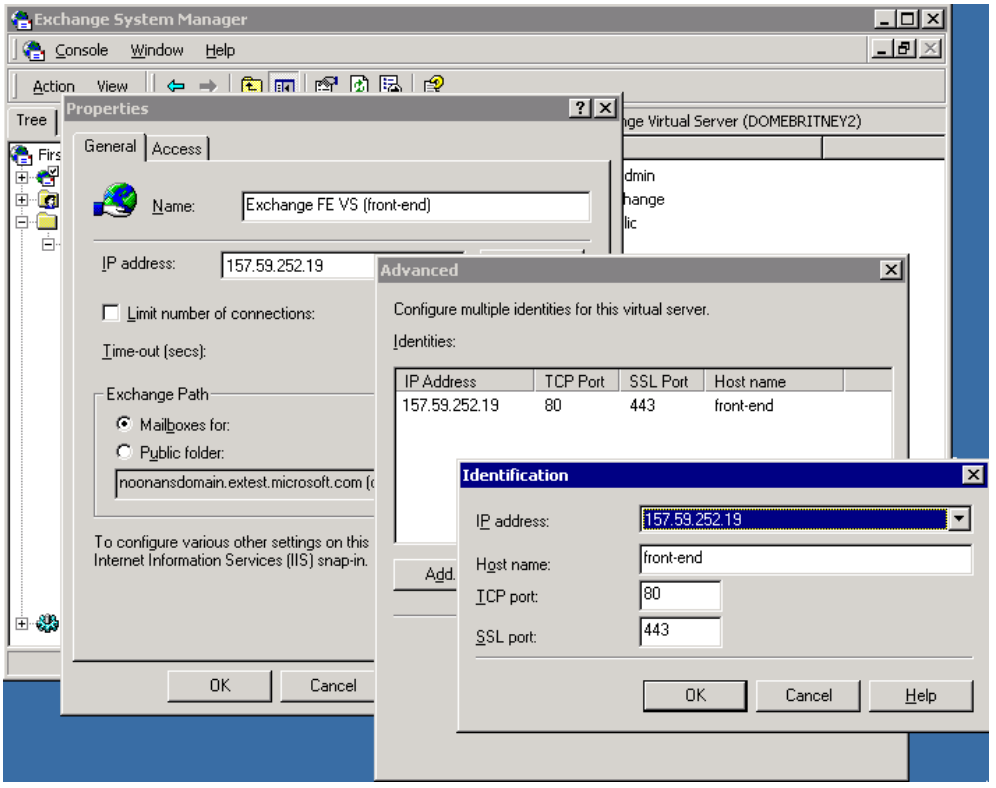
In some circumstances, front-end access to public or private folders might need to be restricted. Remove the "exchange" (for private) or "public" virtual directories from the virtual servers to accomplish this.

The specific back-end configuration steps depend on whether the back-end is hosted on a cluster or not.

Configuring a Non-clustered Back-end

To configure the new virtual server, use the Exchange System Manager, and in the HTTP Protocols container for the specific back-end server, add a new virtual server (right-click on HTTP, select New, Virtual Server...):

- For a name, Microsoft recommends something of the form “Exchange FE VS (*front-end-hostname*)”. Keeping the name consistent will ensure that the purpose of the virtual server can be determined easily.
- Click the **Advanced** button and add an identity for the virtual server that defines, as a hostname, the front-end namespace. This will allow the front-end to communicate correctly with the back-end. You may want to add several hostname entries to the virtual server that list all the ways that a user might access the front-end server. For instance, if a front-end will be used both internally and externally, you may want to list both a short (NetBIOS) name and a fully qualified domain name. See the section below on reducing virtual server creation.
- The TCP port must be port 80, since this is the only port used by the front-end to communicate with the back-end, regardless of the port used by the client to communicate with the front-end. The SSL port may be left at the default.



Configuring the Virtual Server on the Back-end machine

After creating the virtual server, add virtual directories to match the front-end. In most cases, an ”exchange” and a public” virtual directory must be created. Make sure the SMTP domain for the mailboxes matches the front-end’s configured SMTP domain. Do not create an “exadmin” virtual directory on the virtual server. This is only used by the administration tool on the default virtual server, and is not proxied by the front-end.

One Exchange virtual server must be created for each front-end namespace. Thus, in a hosting environment, where many front-end namespaces (xxx.com, yyy.com, etc) are used, there must be a new Exchange virtual server for each of these. This allows for maximum flexibility in determining which resources are available to each hosted company. There need not be a virtual server for every front-end, if they are being used in a load-balancing scenario.

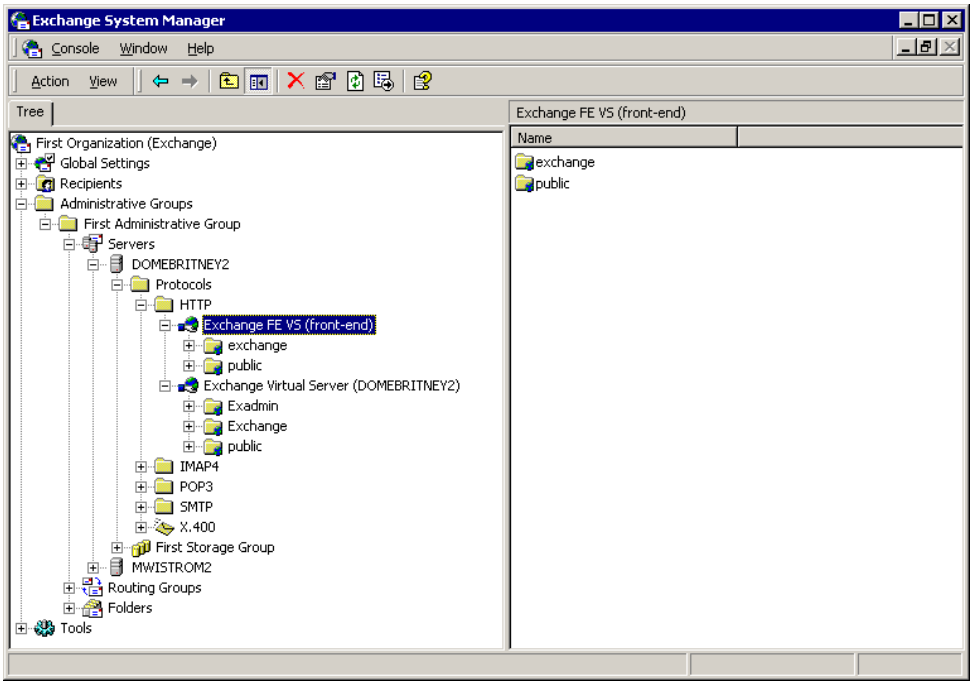


Figure 1: A correctly configured back-end server (assuming a default configuration on the front-end)

In addition, new virtual servers created by the administrator for other reasons (hosting applications, for instance) must be created with the same configuration on the back-ends (i.e. hostnames must the same). *Note*: Virtual servers that differ from the Default virtual server by port number (normal, or SSL) are not supported. All traffic between the front-end and back-end servers is over port 80, regardless of the incoming port used.

Reducing Virtual Server Creation

In some circumstances it may be important to reduce the number of virtual servers created on the back-ends (for instance, if there are thousands of virtual servers). This should not be done unless the administrator fully understands how HTTP virtual servers work. However, if desired, then virtual server creation can be reduced using two methods.

First, if each virtual server on the front-end is configured identically, then a single virtual server can be created each back-end with several hostnames associated with it. In addition, virtual servers need only be created on back-ends which may be directed to by the front-end. For instance, if a front-end contains virtual servers for xxx.com, and for yyy.com, and all xxx.com users are hosted on server A, then on server A, only a virtual server for xxx.com is required. If configurations change, then changes may need to be made to the back-end.

Network Load Balancing (NLB) is an optional availability feature included in Windows 2000 Advanced Server and Windows 2000 Data Center. It enhances availability by distributing client requests among servers, or nodes, participating in a Network Load Balancing cluster. The cluster nodes emit a *heartbeat* every second to announce their availability and when a node fails, the surviving nodes undergo a process called *convergence*, which redistributes the workload among the surviving nodes. NLB clusters should not be confused with Microsoft Cluster Services.

NLB clusters may include up to 32 servers, or nodes. An NLB cluster may be comprised of dissimilar servers, but NLB will balance loads across the cluster regardless of each node’s resources. In such cases, it is possible to set a specific load weight percentage for each cluster node.

Namespace Planning

As mentioned earlier, the FE/BE server topology provides for the use of a single server name or URL to messaging clients. If presenting a single namespace for multiple servers is a requirement, then at least one front-end server will be required. For redundancy and additional client load capacity, many organizations will deploy more than one front-end server. To maintain the single namespace across multiple front-end servers, NLB is recommended. The NLB cluster is assigned a fully qualified domain name (FQDN) such as exchange.company.com, which will be supported by all of the front-end servers, or nodes, in the cluster.

If a single namespace is not required, then messaging clients may

connect to any mailbox resource through any front-end server name, providing that front-end exists in the same Exchange organization. Servers which are not front-ends will support redirection but will not provide single namespace.

URL for OWA

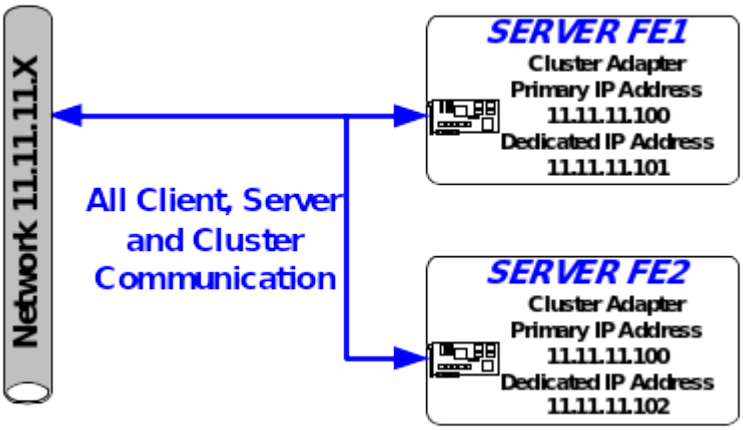
When NLB is configured, the servers or nodes in the cluster are accessible through a name assigned to the cluster. This name will be used for the configuration of messaging clients and, in the case of OWA, as the browser URL. For this reason, select an appropriate and flexible cluster name. While it is possible to change the cluster name without reinstalling the front-end servers, doing so will prevent existing clients from connecting to the (virtual) server.

NLB Configuration

The chief design options around NLB are unicast vs. multicast mode and single NIC vs. Dual NIC. While the choice between unicast and multicast will largely depend on the switching equipment used, the choice between single NIC and dual NIC configurations is often more obvious. A dual NIC configuration allows physical separation of certain client and server communications and assigns them to different network interfaces. This enhances performance as well as performance monitoring since network interface statistics can be tied to the different types of communication.

Single Network Adapter Configuration

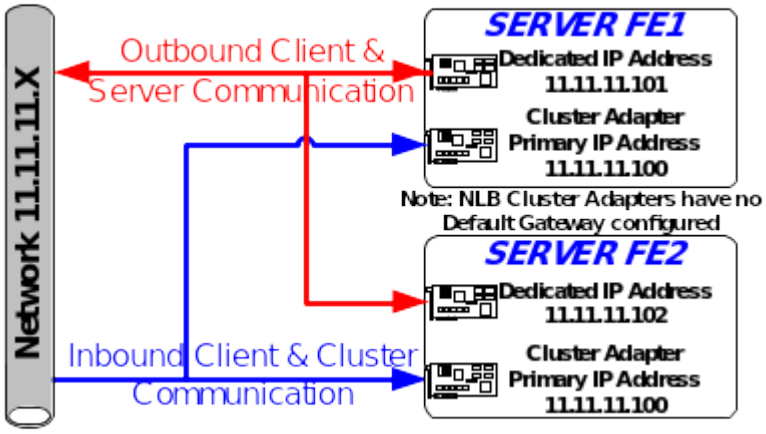
In a single adapter configuration, each front-end server has one adapter that acts as two. As a cluster adapter, each physical adapter receives inbound client requests destined for the cluster's primary IP address. Intra-cluster communication is also handled by the cluster adapter. As a dedicated adapter, each also receives communication via the dedicated IP address, which is intended for that specific server.



Dual Network Adapter (multi-homed) Configuration

The following dual adapter configuration has several advantages over the single NIC configuration:

- Client and Server communication are handled by separate interfaces
- Upstream communication can be handled by separate hubs or switches



The configuration above is enabled through the following settings:

1. Bind NLB Adapter to one NIC on each front-end server.
2. Bind only the primary IP address to the NLB adapter.
3. Remove the Default Gateway for the cluster adapter. This effectively disables all outbound client and server communication from interface.
4. Under NLB properties, the Dedicated IP address can be left blank as all dedicated server communication will be received by the other (dedicated) adapter.
5. Bind each server's dedicated IP address to the other NIC.

Register the cluster FQDN in DNS. Make sure to register in external DNS, too, if the cluster FQDN is to be used by Internet messaging clients.

NLB Configuration Parameters

Client Affinity settings assist in preserving client session states by mapping clients to the same cluster node. This is particularly important for SSL sessions, which are maintained between a client and a unique host. Using the default "Single" client affinity best supports SSL sessions as it causes NLB to examine the client's full IP address to maintain affinity.

Testing NLB

In order to test Network Load Balancing, you need to Ping the fully

qualified NLB cluster name and IP from client network (internal or external).

If you are using the default unicast load balancing mode, the cluster nodes cannot communicate with each other unless they have second network adapters.

The WLBS.exe command can be used to test NLB arrays.

More Information

For further information on NLB, see the following references:

http://support.microsoft.com/support/kb/articles/Q193/6/02.ASP

Exchange 2000 Resource Kit, Chapter 25: Outlook Web Access.

Basic Authentication is the only authentication method supported in a FE/BE server topology. To prevent clear text user names and passwords from being intercepted, Secure Sockets Layer (SSL) encryption must be used to provide a secure channel for logon and messaging.

Configuring SSL requires the installation of a Web Server certificate. Organizations requiring SSL for an Intranet deployment may install Windows 2000 Certificate Server as an Enterprise Certificate Authority. Organizations enabling secure Internet solutions can purchase server certificates from Certificate Authorities such as Verisign.

For information on obtaining and installing server certificates, please refer to Step-by-Step Guide to Internet Protocol Security (IPSec)

http://www.microsoft.com/WINDOWS2000/library/planning/security/ipsecsteps.asp

Install server certificate on every front-end server

Once a server certificate has been correctly installed, encryption is enabled for the registered server name. Test Outlook Web Access using an SSL-encrypted session prior to requiring it for all users. Invoke Secure HTTP by using **https://** in front of the OWA URL.

If all users will be required to use SSL-encrypted sessions, perform the following procedure on each Front-End server..

Steps:

1. In Internet Services Manager, right-click the Web site (e.g., Default Web Site) or the virtual directory (e.g., Exchange) to which you want to require secure access and select **Properties**.
2. Select the Directory Tab.

OPERATIONAL RECOMMENDATIONS

3. Select **Require Secure Channel when accessing this resource**.
4. Click **Apply**, click **OK**, and then click **OK** again.
5. You may need to stop and restart the Web site.

Remember that SSL encryption/decryption requires front-end server processor overhead and carries an estimated performance penalty of 15 percent or more.

Disabling Front-End Services

Once promoted, front-end servers need only run minimal Exchange services. In the case of front-end servers supporting only HTTP (OWA) clients, all Exchange services can be stopped. Stopping and disabling all unnecessary services is recommended to reduce the number of processes on the front-end server and to harden the server against attacks if it resides in an Internet-exposed perimeter network.

For each protocol supported, configure Exchange services to run as follows:

- **HTTP**: Stop all Exchange-related services and set for manual start up.
- **POP3**: Stop all Exchange-related services except Exchange Pop3, Exchange System Attendant and Exchange IS. Other Exchange services should be set for manual start up.
- **IMAP4**: Stop all Exchange-related services EXCEPT Exchange IMAP4, Exchange System Attendant and Exchange IS. Other Exchange services should be set for manual start up.

To further reduce the number of required services for front-end servers, the Exchange POP3 and Exchange IMAP4 services can both be configured to run without the Exchange IS. To disable the dependency of these services on Exchange IS, refer to the Registry configuration information in the MS white paper, “Exchange Front-End Back-End Server Topology” (see Other Resources).

Setting up a Front-end/Back-end Topology: Checklist

Here is a quick checklist of what you need to configure.

- ☑ Install the front-end server
  - o Remove public and private stores from the server
  - o Stop and disable all unnecessary services
  - o *Optional:* Enable SSL encryption

- *Optional:* Assign front-end IPSec Policy
- *Optional:* Set up hosts file
- *Optional:* Set up directory access registry entries

☑ Configure the back-end servers

- Create virtual servers to match the front-end.
- Create virtual directories to match those on the front-end.
- *Optional:* Assign back-end IPSec policy

☑ *Optional:* Configure firewalls

- Enable access to protocol ports

☑ *Optional: Enable access to additional ports.*

Backup

Let's briefly consider the requirements for backing up front-end/back-end configurations.

FE Backup Requirements

Organizations will need to determine the backup requirements of their front-end servers based upon their service level goals. Since front-end servers run minimal services and do not typically host any client data, they will not need to be recovered as rapidly as back-end servers in order to meet availability requirements. Even so, front-ends do provide for client connectivity and could reduce application responsiveness if unavailable. Therefore, the following should be considered when determining the appropriate backup policies for front-end servers.

- Can the front-end server be re-installed faster than it can be restored?
- Based upon the number of front-end servers in your environment, what would the impact to client connectivity be if a front-end failure occurred during a peak period?

You may find it more cost effective to just have more than one Front-end server, accept the reduced performance if one goes down and rebuild it again from a CD image if required.

Modifying the Default Exchange Virtual Directories

A default virtual directory is created automatically for the virtual server and a control virtual directory is created for all newsgroups. You can create a new virtual directory for each newsgroup or you can modify the default directory. When you modify a directory, you can change the path to the public folder in which messages are stored, secure access to the virtual directory, restrict access to the virtual directory, and control content.

To modify a virtual directory:

1. Start System Manager: On the Start menu, point to **Programs**, point to **Microsoft Exchange**, and then click **System Manager**.

Cannot Start the Application

You can manually start the application by following these instructions: On the **Start** menu, point to **Programs**, point to **Microsoft Exchange**, and then click **System Manager**.

2. Navigate to **Virtual Directories**.
    1. *Server*
    2. Protocols
    3. NNTP
    4. *NNTP virtual server*
    5. Virtual Directories
3. Right-click a virtual directory, and then click **Properties**.
4. To change the public folder in which messages are stored, click **Contents**. Select **Directory located on this computer**, **Share located on another computer**, or **Custom resource**, and then type the path to the folder or to the network directory, user name, and password.
5. To require that all access to the directory takes place on a secure channel, in **Properties**, click **Secure**. To use encryption if a certificate is installed, select **Require secure channel**. If the computer running Exchange supports 128-bit encryption, click **Require 128-bit Encryption**.
6. To allow clients to post to a newsgroup's virtual directory, in **Properties**, select **Allow posting** if it is not already selected.
7. To list this newsgroup only if the user has the same ACL permissions as the newsgroup folder, select **Restrict newsgroup visibility**. If you select this option and do not set ACL permissions on the newsgroup folder and the user accounts, this newsgroup will not be available to users.
8. To log entries every time a virtual directory is accessed, select **Log access**.
9. To create an index of news articles, select **Index news content**.

SUMMARY AND NEXT STEPS

In this chapter you looked at the configuration steps necessary to install and operate Front-end/Back-End messaging servers using Exchange 20000. Messaging environments of all sizes that support Internet standard clients can realize scalability and availability improvements through the deployment of this Exchange 2000 FE/BE server topology. As a multi-tiered application architecture, Exchange 2000 provides a strategic advantage to growing organizations that also need to tailor messaging service levels on a per-protocol basis.

After you have deployed a FE/BE server topology, there are still further design and operations components to consider. Depending on your organization's application scaling requirements, some suggested next steps include:

Hardware load balancing

Dedicated hardware load balancing solutions such as f5's Big IP and Cisco's LocalDirector can provide greater load balancing options and scalability than Windows 2000 NLB. These solutions can balance loads across unlimited hosts and intelligently monitor application data to ensure server availability.

SSL acceleration

Dedicated SSL accelerator cards from manufacturers such as Rainbow Technolgies and nCipher may be installed in front-end servers or hardware load balancing appliances to eliminate the processing overhead associated with SSL encryption and decryption algorithms.

More Information

For more information on Outlook Web Access, see the following link:

http://www.microsoft.com/TechNet/exchange/e2kowa.asp